An incipient fault diagnosis methodology using local Mahalanobis distance: Detection process based on empirical probability density estimation

Junjie Yang, Claude Delpha

Version of Record: https://www.sciencedirect.com/science/article/pii/S0165168421003455
Manuscript_cfecf8f2cbb9ff935a520c2b7965669b

An Incipient Fault diagnosis Methodology Using Local Mahalanobis Distance: Detection process based on Empirical Probability Density Estimation

Junjie Yang, Claude Delpha*

Universite Paris Saclay, CNRS, CentraleSupelec, Laboratoire des Signaux et Systemes (L2S), 3 Rue Joliot Curie, Gif sur Yvette, France

ARTICLE INFO

Keywords:
Incipient fault diagnosis
Fault Detection
Time varying fault
Mahalanobis distance
Multivariate statistical process monitoring

Abstract

Incipient fault detection is growing as a challenging and hot topic in industrial and academic areas. It is essential to avoid slight unpermitted changes of a system state that can be aggravated and lead to severe security issues. The main challenge of this problem lies in the fact that tiny changes in the early stage can be blurred with noise and create confusion leading to poor detection performance of typical fault detection methods. To detect subtle deviations buried in noise and cope with the non-Gaussian distributed data condition while keeping with the time series information, a sensitive fault detection methodology combining a specifically tuned Local Mahalanobis Distance (LMD) algorithm and an Empirical Probability Density (EPD) estimation technique is proposed. More specifically, first, a healthy domain estimation is proposed to compute the local Mahalanobis distance with optimally tuned characteristics. To approximate a healthy domain, this work proposes a down-sampling algorithm for anchors generation and a parameter estimation method optimally tuned and based on Generalized Extreme Value distribution (GEV) for the domain margin selection. Subsequently, the EPD cumulative sum technique is applied to the LMD result for improving the detection sensitivity further. The performance analysis based on simulation data shows that our proposal is effective to non-Gaussian data and sensitive for incipient fault detection. A case study based on the Continuous-flow Stirred Tank Reactor (CSTR) further validates the effectiveness of our proposal and highlights its benefit by comparing it with state-of-the-art-based solutions in terms of detection delay, detection probability, false alarm probability, and area under the receiver operating characteristic curve (AUC).

1. Introduction

Health monitoring plays a key role in modern industrial systems for increasing requirements on reliability, availability, maintainability, and safety. With the development of sensor technology and computational science, fault diagnosis, including fault detection, faulty variables isolation and fault severity estimation, has become a more and more hot topic in the last decade. Its applications cover abundant scenarios such as chemical manufacturing, electrical system, and mechanical process [1, 33, 39].

Generally, a fault, defined as "*the unpermitted deviation of at least one process parameter from an acceptable condition*", can lead to system performance degradation or even an unexpected stop [11, 29]. To avoid a severe accident and reduce economic impact, protection actions or necessary fault-tolerant control are required for systems with faults, which highly rely on quick and accurate fault detection. Therefore, it is essential to develop effective fault detection methodologies with a short detection time delay.

Motivated by the above goal, masses of studies were presented in the literature to characterize fault behavior and provide solutions for fault detection. Typical methods like Shewhart charts, cumulative sum, and EWMA (Exponentially Weighted Moving Average) provided a simple and effective univariate measurement to characterize process behavior [27, 22, 17, 25]. Then, parts of these methods were extended for multivariate analysis, such as MEWMA (Multivariate Exponentially Weighted Moving Average) and the MCUSUM (Multivariate Cumulative Sum) [18, 8].

For high dimensional data, multivariate statistical techniques were proposed to reduce the data dimension and further improve the detection capability by considering the relevant information lying on a low dimensional subspace. As an example, fault detection approaches using PCA (Principal Component Analysis), ICA (Independent Component Analysis), and FDA (Fisher Discriminant Analysis) have been well studied and compared in [34]. In the multivariate statistical framework, Hotelling's T^2 and SPE(Q) statistics are widely used as indexes to determine if observed data is different from the healthy one [35]. However, one drawback of these methods lies in their ineffectiveness for non-Gaussian data or non-linear processes [34].

To cope with this problem, methods focusing on the relation of system input and output were proposed, e.g., PLS (Partial Least Squares) [19], GCCA (Generalized Canonical Correlation Analysis) [4], CVA (Canonical Variate Analysis) [21], and CVDA (Canonical Variate Dissimilarity Analysis) [24]. Despite the pronounced performance of existing approaches, fault detection for non-Gaussian distributed data and non-linear systems with minimal false alarm and missed detection remains an open issue.

Recently, intelligent approaches based on machine learning theories are widely used in fault detection problems for their powerful feature representation capability. In particular, neural network-based methods, which can automatically extract faulty features and recognize different kinds of samples, are attracting much interest from academics [30,

*Corresponding author
ORCID(s): 0000-0003-3224-8628 (C. Delpha)

5, 15]. However, major neural networks methods are supervised and require a large number of labeled data for both healthy and faulty scenarios, which is impractical for industrial monitoring since faulty behaviors are diverse and not always predictable. The lack of faulty knowledge induced the One-Class Classification (OCC) problem, for which the auto-encoders method, one-class Support Vector Machine (OSVM), and k-centers are common solutions [31, 12, 37]. Their drawback is either the weak robustness to noise disturbances or the requirement of enormous sample data set.

Furthermore, another hot challenge of the fault detection problem is that subtle deviation is usually buried in noise, leading to the difficulty of detection. The subtle deviation is the so-called incipient fault described as a slowly varying change of parameters or a disturbance with tiny fault amplitude compared to the noise power [11, 10]. Unfortunately, most of the aforementioned approaches are reported as not sensitive enough to detect incipient faults except the CVDA method [24].

In contrast to multivariate statistical techniques, some divergence-based approaches introduced in [38, 6, 9, 40] employ the Kernel Density Estimation (KDE) method to estimate the probability distribution and evaluate the dissimilarity of observed data and healthy data. Although these methods show the benefit of using the Kullback–Leibler and Jensen–Shannon divergence in incipient fault detection, they seem unsuitable for online applications due to the strong dependence on a large sample number to estimate probability distribution.

Recently, a method combining the Empirical Probability Density (EPD) and Cumulative sum (CUSUM) techniques is proposed in [32] to provide a sensitive solution for incipient fault detection. However, the so-called EPD-CUSUM method assumes that a reference fault-free signal with the same time-scale as the faulty one is available, which is impractical for most industrial applications.

Motivated by the above discussion, this work proposes an incipient fault detection methodology with the goals of 1) coping with the non-Gaussian and non-linear cases, 2) without using prior knowledge of the faulty behavior, 3) improving detection sensitivity. The contributions associated with our proposal are composed of the following three points.

1. Considering samples' spatial distribution, this work first proposes a method to determine the healthy domain of fault-free samples with multiple anchors and a domain margin. Based on the approximated healthy domain, a key index named the local Mahalanobis distance (LMD), is defined as the closeness measurement of observed samples concerning the healthy domain. Although the Mahalanobis distance (MD) is used to measure the closeness of samples and a Gaussian distribution, the goal of our LMD proposal is to keep the advantages of MD and cope with the ineffectiveness for non-Gaussian cases. Then this allows to reach a good efficiency for non-linear systems. The properties of the LMD, including its probability density distribution and the behavior in healthy conditions, are also studied and introduced to provide a theoretical foundation for fault detection.
2. In the healthy domain approximation procedure, an optimized down-sampling algorithm for anchors generation, extracting the spatial distribution information of the healthy domain, is developed. The optimization of this procedure is performed to make a trade-off between computation cost and approximation error. Additionally, for the selection of the healthy domain margin, a convex optimization problem based on the generalized extreme value distribution (GEV) as the distribution model of LMD is established and solved.
3. Finally, our early proposed EPD-CUSUM method is improved to overcome its primary drawback and become applicable to the LMD index. Then, it is specifically combined with the LMD index as the proposed incipient fault detection framework to increase sensitivity further. The benefit of using the improved EPD-CUSUM method is shown by comparing it with the original LMD index and the typical cumulative sum method, which again validates its excellent incipient detection capability.

The remaining part of this paper is organized as follows. Section 2 formulates the fault detection problem and introduces the challenge in the non-Gaussian distribution case. The healthy domain approximation for fault detection is presented in section 3, including the definition of LMD, the down-sampling algorithm for anchors generation, and parameters estimation of the GEV model for margin selection. Then the fault detection framework based on LMD and EPD-CUSUM is proposed and detailed in section 4. Subsequently, the performance of our proposal is evaluated by considering different noise levels and fault severity in section 5. In section 6, the CSTR process is introduced as a case study for the effectiveness validation, and the performance of the proposed method is compared with existing methods. Finally, section 7 concludes this work.

2. Problem formulation and background

Let's consider a sample vector $\boldsymbol{x} \in \mathbb{R}^m$ with m variables representing features or original signals, for example, electrical signal, temperature, pressure, and concentration. In the fault diagnosis context, it is natural to decide $\boldsymbol{x}$ as being healthy or faulty. From the view of spatial distribution, $\boldsymbol{x}$ belongs to one of the two domains in m dimension space, such as the healthy domain denoted as $\mathcal{H}$ and the faulty one denoted as $\mathcal{F}$, where $\mathcal{H} \cap \mathcal{F} = \emptyset$ and $\mathcal{H} \cup \mathcal{F} = \mathbb{R}^m$. From this point of view, fault detection turns to determine such a separating hyperplane $\mathcal{B}(\boldsymbol{x}) - b = 0$ satisfying

$$\mathcal{B}(\boldsymbol{x}) - b \leq 0, \text{ for } \boldsymbol{x} \in \mathcal{H}$$
$$\mathcal{B}(\boldsymbol{x}) - b > 0, \text{ for } \boldsymbol{x} \in \mathcal{F}$$

where $\mathcal{B}(\cdot) : \mathbb{R}^m \to \mathbb{R}^1$ and b is an offset.

Although supervised approaches, such as Support Vector Machines (SVM), can effectively achieve an optimal sepa-

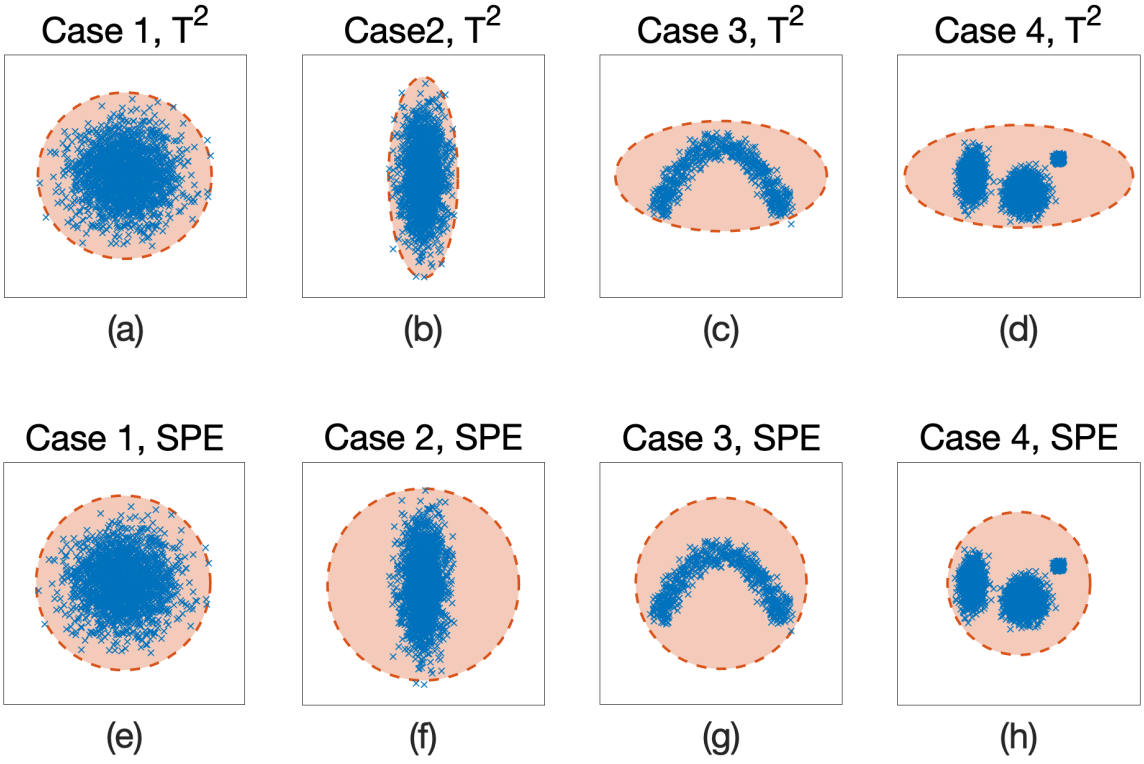

Figure 1: Healthy domain approximation based on Hotelling's T^2 and SPE(Q) statistics for different cases. Case 1: two Gaussian distributions with same variance; Case 2: two Gaussian distributions with different variances; Case 3: two sine functions; Case 4: Gaussian mixture distribution.

ration based on negative and positive samples [16], it is impractical for industrial applications. The lack of sufficient faulty samples in the training step induces the more challenging OCC problem. A primary solution for the OCC problem is to obtain the description of the healthy domain in feature space based on historical fault-free samples, which is called healthy domain approximation.

For example, hypersphere in m dimension space is commonly used as a model to approximate the healthy domain $\mathcal{H}$. Two common PCA-based statistics, namely Hotelling's T^2 and SPE(Q), can be regarded as two examples of healthy domain approximation, where Hotelling's T^2 is the particular case of the SPE(Q) statistics considered in normalized principal component space. The two statistics can be expressed in the form of separating hyperplane as follows

- Hotelling's T^2: $\mathcal{B}(\boldsymbol{x}_{pc}) = \boldsymbol{x}_{pc}^T \Lambda^{-1} \boldsymbol{x}_{pc}$, $b = g_{th,T^2}$
- SPE(Q): $\mathcal{B}(\boldsymbol{x}_{res}) = \boldsymbol{x}_{res}^T \boldsymbol{x}_{res}$, $b = g_{th,spe}$

where $\boldsymbol{x}_{pc}$ and $\boldsymbol{x}_{res}$ are transformed data with zero mean in L dimension principal space and residual space, respectively; $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_j, \cdots, \lambda_L)$ is the corresponding eigenvalues matrix whose element λ_j is also the variance of the j^{th} variable of $\boldsymbol{x}_{pc}$; g_{th,T^2} and $g_{th,spe}$ are thresholds for these two statistics.

However, Hotelling's T^2 and SPE(Q) statistics are only effective to elliptically distributed data, e.g., for multivariate Gaussian distribution and multivariate t-distribution. As shown in Fig.1, SPE(Q) statistics is only effective in case 1, the Gaussian distribution with the same variance. For case 2, the two Gaussian distributions with different variances, the approximate healthy domain determined using SPE(Q) is dramatically larger than the actual distribution, see Fig.1-(f). Although for both cases 1 and 2, Hotelling's T^2 can fit data perfectly for its normalization of each dimension, it can not specify an accurate border of the healthy domain for non-elliptically distributed data, see cases 3 and 4.

In fact, signals like cases 3 and 4 are quite common in industrial applications, but most fault detection approaches are only dedicated to Gaussian distribution cases. Therefore, this work aims at finding an accurate description of the healthy domain for non-elliptically distributed data by only using fault-free samples.

3. Healthy domain approximation

Let the fault-free sample matrix be $\boldsymbol{X}^*$. To specify the healthy domain's border, we focus on the samples' spatial distribution feature while ignoring their probability density distribution. For the irregularity of spatial distribution, we use multiple hyperspheres with the same radius but different

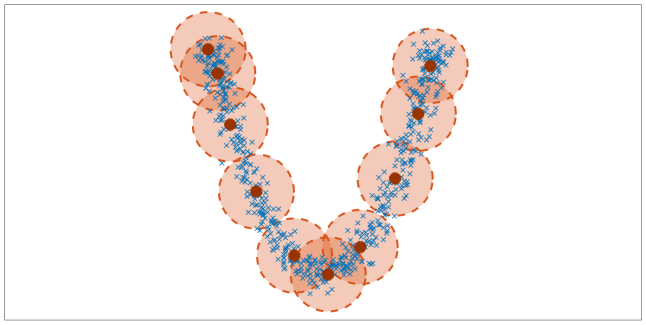

Figure 2: Proposed idea for healthy domain approximation (red points are the centers of hyperspheres)

centers to describe the healthy domain.

As illustrated in Fig.2, with the position of centers and the radius's value being properly specified, the overlapping hyperspheres can perfectly cover all the fault-free samples. These centers, denoted as $\boldsymbol{C}$ and referred to as anchors hereafter, contain essential spatial information of the healthy domain. Likewise, the radius denoted as ϵ specifies the margin of the domain. In other words, an approximate healthy domain is completely determined by an anchor set S such as $S = \{\boldsymbol{C}_1, \cdots, \boldsymbol{C}_k, \cdots\}$, $S \subseteq \mathcal{H}$ and the healthy domain margin ϵ. Consequently, a fault detection procedure can be performed by deciding if a sample is inside the healthy domain.

To obtain such an approximate description of $\mathcal{H}$, a particular distance measure is first defined, providing mathematical tools for the later analysis. Approaches for anchors generation and domain margin selection will be introduced in the rest of this section.

3.1. Local Mahalanobis distance

In the concept of healthy domain approximation, distance measure [3] plays a central role, motivating the following discussion on the distance measure.

In multivariate applications, Mahalanobis distance attracts more interest than Euclidean distance because the former is unitless and scale-invariant and considers the correlations of variables. Suppose $\boldsymbol{A}, \boldsymbol{B}$ are two sample vectors generated from the same distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. Then the multivariate Mahalanobis distance of $\boldsymbol{A}$ with respect to $\boldsymbol{\mu}$ is defined as:

$$d_M(\boldsymbol{A}, \boldsymbol{\mu}) = \sqrt{(\boldsymbol{A} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{A} - \boldsymbol{\mu})} \tag{1}$$

The Mahalanobis distance between $\boldsymbol{A}$ and $\boldsymbol{B}$ is

$$d_M(\boldsymbol{A}, \boldsymbol{B}) = \sqrt{(\boldsymbol{A} - \boldsymbol{B})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{A} - \boldsymbol{B})} \tag{2}$$

Based on Mahalanobis distance, the local Mahalanobis distance (LMD), denoted as D, is defined to indicate the distance between a sample $\boldsymbol{x}$ and $\mathcal{H}$, such as:

$$D(\boldsymbol{x}; \mathcal{H}) \overset{def}{=} \min_k \left\{ d_M\left(\boldsymbol{x}, \boldsymbol{C}_k\right) | \boldsymbol{C}_k \in \mathcal{H} \right\} \tag{3}$$

Geometrically, $D(\boldsymbol{x}; \mathcal{H})$ is the closest Mahalanobis distance from a sample to the domain $\mathcal{H}$. In the healthy domain approximation procedure, $\mathcal{H}$ is replaced by S to obtain the estimated distance $D(\boldsymbol{x}; S)$ which is denoted as ξ and calculated as:

$$\xi = D(\boldsymbol{x}; S) = \min_k \left\{ d_M\left(\boldsymbol{x}, \boldsymbol{C}_k\right) | \boldsymbol{C}_k \in S \right\} \tag{4}$$

Let κ be the element number of S. According to Eq.(4), three properties of ξ are summarized as follows:

1. $\xi \in [0, +\infty)$
2. For healthy samples ($\boldsymbol{x} \in \mathcal{H}$), $\lim_{\kappa \to +\infty} \xi = 0$. Since $\boldsymbol{x} \in \mathcal{H}$ and $S = \mathcal{H}$ when $\kappa \to +\infty$, we have

   $$\xi = D(\boldsymbol{x}; S) = \min_k \left\{ d_M\left(\boldsymbol{x}, \boldsymbol{C}_k\right) | \boldsymbol{x}, \boldsymbol{C}_k \in S \right\}$$

   Then, there is always a $\boldsymbol{C}_k$ whose value is equal to $\boldsymbol{x}$, and the minimum value is always 0.
3. According to the theorem of extreme value statistics [2, 13], whose introduction is given in appendix A, the distribution of ξ can be modeled by the following general form named Generalized extreme value distribution (GEV) for minima

   $$\Phi(\xi; \rho, \beta, \tau) = 1 - \exp\left\{ -\left[1 + \tau\left(\frac{\rho - \xi}{\beta}\right)\right]^{-1/\tau} \right\}$$
   $$-\beta - \tau(\rho - \xi) \leq 0, \qquad \beta > 0 \tag{5}$$

   where ρ, β, τ respectively the location, scale and shape parameters.

The three properties of LMD will be used in the following two parts to obtain optimal anchors and determine the domain margin.

3.2. Down-sampling algorithm for anchors generation

As previously mentioned, the healthy domain approximation for our proposal has to be done using a set of anchors. This part aims at generating these anchors based on fault-free samples.

One of the goals of the anchors-generation procedure is to remove redundancy and extract critical spatial information from the original sample set. More precisely, geometrically close samples are identified and merged as an anchor. Although clustering algorithms, like k-means, can achieve a similar goal, they are either easily affected by outliers or cannot be optimized in the proposed framework. Therefore, a distance-based down-sampling algorithm with its optimization procedure is proposed to generate anchors.

Using the original fault-free samples matrix denoted as $X^*_{K \times m} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_K]^T$, where K is the number of samples and m is the number of variables, the procedure is drawn according to the following main steps:

1. Calculate samples' Mahalanobis distance $d_M(\boldsymbol{x}_i, \boldsymbol{\mu})$ considering their mean center $\boldsymbol{\mu}$.

2. For initialization, rearrange these samples as a queue $\mathbb{Q}$ in ascending order according to their d_M value.
3. Then, take one sample $\boldsymbol{x}_i$ from $\mathbb{Q}$ each time as the center of a region whose radius is given by γ and find out all the samples inside this region, denoted as Z, such as:

$$Z_k = \left\{ \boldsymbol{x}_l | d_M\left(\boldsymbol{x}_i, \boldsymbol{x}_l\right) \leq \gamma, \quad \boldsymbol{x}_l \in \mathbb{Q} \right\} \tag{6}$$

4. Afterward, calculate the mean of samples in Z_k to obtain an anchor $\boldsymbol{C}_k$

$$\boldsymbol{C}_k = \bar{Z}_k = \frac{\sum_{l=1}^{\ell_z} \boldsymbol{x}_l}{\ell_z} \tag{7}$$

where ℓ_z is the element number of Z_k.
5. Subsequently, remove elements of Z_k from $\mathbb{Q}$. These samples will be represented by their anchor $\boldsymbol{C}_k$.
6. Repeat steps 3 to 5 until $\mathbb{Q}$ is empty.
7. All the obtained anchors are grouped into the anchors set $\boldsymbol{S}$.

In the down-sampling procedure, the region radius γ is crucial to control the anchors' number κ: κ decreases as γ increases. In fact, if $\gamma \to 0^+$, then $\boldsymbol{S}$ is composed of all the fault-free samples and $\kappa = K$. In the opposite, when $\gamma \to \infty$, then $\boldsymbol{S}$ contains only one element.

According to the definition and properties of LMD, κ is related to computational complexity and approximation error. On the one hand, a large value of κ (γ is small) causes a slow calculation speed since the complexity of LMD computation is $\mathcal{O}(\kappa)$. On the other hand, a large κ helps to reduce approximation error because ξ gets close to 0 in this case. As an example, Fig.3 shows the evolution of ξ and κ according to γ. A slight increase of γ (reducing κ) can lead to acceptable approximation error but keep a reasonable computational cost, which motivates us to trade off these two factors.

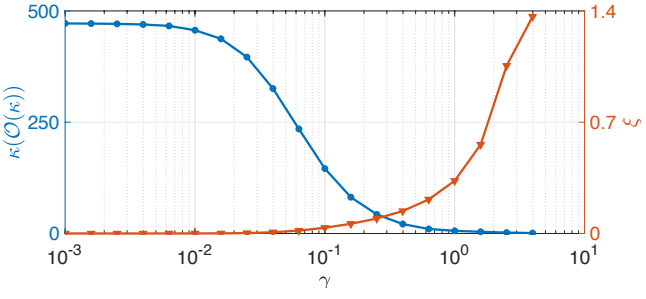

Figure 3: Evolution of LMD and the anchors' number (computational complexity) against different region radiuses

To quantify the approximation error, the mean square error with normalized coefficient η is used as the cost term.

$$\text{Err} = \frac{\frac{1}{K}\sum_{i=1}^{K} D^2(\boldsymbol{x}_i; S)}{\eta} \tag{8}$$

where the normalized coefficient η can be obtained as the maximum error. As discussed above, when $\gamma \to +\infty$, the error becomes maximum, and there is only one anchor, i.e., $\boldsymbol{S} = \{\boldsymbol{\mu}\}$. Thus, η can be calculated as follows:

$$\eta = \frac{\sum_{i=1}^{K} D^2(\boldsymbol{x}_i; \{\boldsymbol{\mu}\})}{K} = \frac{\sum_{i=1}^{K} d_M^2(\boldsymbol{x}_i, \boldsymbol{\mu})}{K} \tag{9}$$

Additionally, the compression rate value Cr is employed to quantify the computational cost since the computational complexity is related to κ. It can be written as:

$$\text{Cr} = \frac{\kappa}{K} \tag{10}$$

Ultimately, the following cost function is proposed and minimized to obtain the optimal radius denoted as γ^{opt}.

$$\text{Loss} = \text{Cr} + \text{Err} \tag{11}$$

$$\gamma^{opt} = \arg\min_{\gamma}(\text{Loss}) \tag{12}$$

Based on this knowledge, the optimization of the loss function can be done to keep the best radius size resulting in the minimum computational cost and approximation error. Thus, the anchors generated from this down-sampling operation are directly used in LMD calculation.

3.3. Parameter estimation of GEV for domain margin ϵ

According to the third property of LMD, once the parameters of the probability distribution model of ξ are estimated, one can select the corresponding domain margin based on the probability model for a given significance level α. This approach helps to obtain a robust decision boundary for healthy domain approximation by mitigating the influence of outliers.

To that end, we minimize the mean squared error of the estimated general extreme value distribution model and the empirical cumulative density. Then the parameters of the model are determined when the error reaches the minimum. This parameter estimation procedure can be described by the following convex optimization problem.

$$\begin{aligned} &\underset{\rho,\beta,\tau}{\text{minimize}} \ \frac{1}{K}\sum_{i=1}^{K}\left[\Phi\left(\xi_i;\rho,\beta,\tau\right) - \mathcal{P}(\xi_i)\right]^2 \\ &\text{subject to} \ -\beta - \tau(\rho - \xi_i) \leq 0, \qquad \beta > 0 \end{aligned} \tag{13}$$

where $\Phi(\xi_i;\rho,\beta,\tau)$ is the GEV distribution function, $\mathcal{P}(\cdot)$ is the empirical cumulative density. To solve the convex optimization problem with multiple constraints, we employ the Nelder–Mead method to search solutions iteratively until the objective function (Eq.13) converges [20]. Nelder–Mead method is a numerical technique allowing to search minimum or maximum in a multidimensional space. It often serves as a basic tool for convex optimization problems when derivatives of the objective function are unknown or difficult to calculate [20, 28, 14]. After determining the model's parameters, the healthy domain margin ϵ is determined with

the given significance level α

$$\epsilon = \Phi^{-1}(\alpha;\rho,\beta,\tau) = \rho + \frac{\beta}{\tau} - \frac{\beta\left[-\ln(1-\alpha)\right]^{-\tau}}{\tau} \quad (14)$$

As examples, the proposed healthy domain approximation method is applied to the four cases previously presented in section 2. The obtained results are displayed in Fig.4. The results show that this approach yields a more representative healthy domain approximation than the classical Hotelling's T^2 and SPE(Q) methods: it covers most of the fault-free samples but excludes outliers.

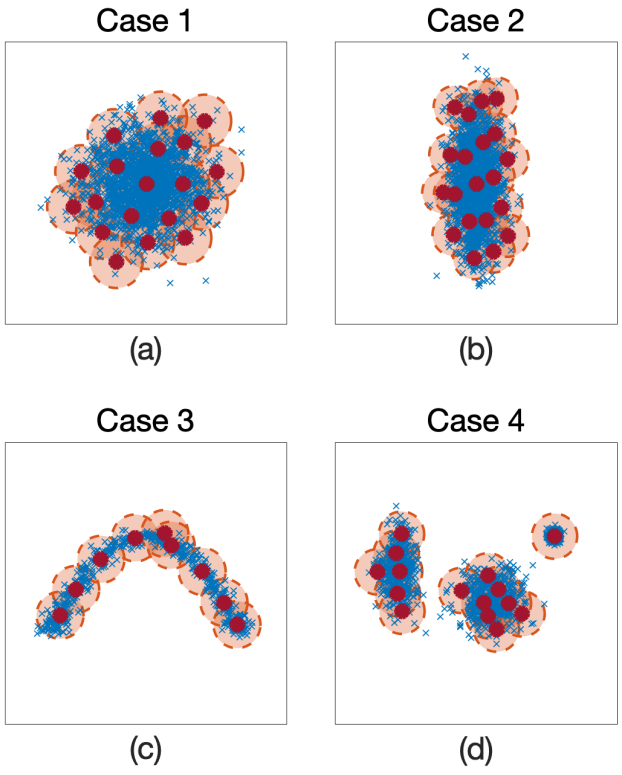

Figure 4: Healthy domain approximation based on the proposed method for different cases

4. Fault detection

Our proposal for incipient fault detection, referred to as Local Mahalanobis Distance Analysis (LMDA), can then be presented in two parts: the training process and the online monitoring. The main steps of these two parts are summarized in Fig.5 and described in the following subsections.

In order to go through the explanations for the proposal and show a step-by-step validation, we consider, in the following simulations, a multivariate data system example referred from the paper [36].

The considered eight-variable system with K values is:

$$\begin{aligned} x_1(t) &= 1+\sin(0.1t), & x_2(t) &= 2\cos^3(0.25t)\cdot e^{-t/K} \\ x_3(t) &= \log(x_2(t)^2), & x_4(t) &= x_1(t)+x_2(t) \\ x_5(t) &= x_1(t)-x_2(t), & x_6(t) &= 2x_1(t)+x_2(t) \\ x_7(t) &= x_1(t)+x_3(t), & x_8(t) &\sim \mathcal{N}(0,1) \end{aligned} \quad (15)$$

In the following, let the signal vector $\boldsymbol{x}^* = (x_1, \cdots, x_8)$ stands for a fault-free and noise-free sample case. The healthy and faulty sample vectors respectively denoted as $\boldsymbol{x}^h$ and the $\boldsymbol{x}^f$ can be described as:

$$\boldsymbol{x}^h = \boldsymbol{x}^* + \boldsymbol{v} \quad (16)$$

$$\boldsymbol{x}^f = \boldsymbol{x}^* + \boldsymbol{f} + \boldsymbol{v}' \quad (17)$$

where $\boldsymbol{f}$ is the fault component, $\boldsymbol{v}$ and $\boldsymbol{v}'$ are noise components whose signal-to-noise ratio (SNR) is settled to 20dB as a default value. Without loss of generality, for this work, the fault is modeled as a single and time-varying change occurring at variable c, such as $\boldsymbol{f} = [0, \ldots, t\delta, \cdots, 0]$, where δ is a constant linked to the fault severity.

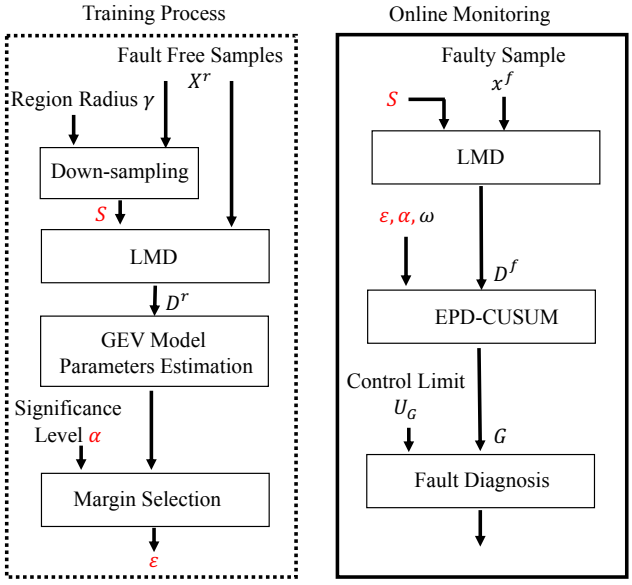

Figure 5: Proposed LMDA-based diagnosis procedure

4.1. Training process

The goal of this training process is to obtain the optimal anchors set S and determine the domain margin ϵ to provide a setting for online monitoring. In this part, we consider a fault-free data matrix $\boldsymbol{X}^h$ for training with $K \times m$ dimensions such as $\boldsymbol{X}^h_{K\times m} = [\boldsymbol{X}_1, \cdots, \boldsymbol{X}_i, \cdots, \boldsymbol{X}_K]^T$.

The first step consists in performing the down-sampling algorithm on $\boldsymbol{X}^h$ with different region radius γ to search for the optimal one. The loss function given in Eq.(12) is calculated and its results are shown in Fig.6. When the loss reaches its minimum value, the searching procedure stop and the corresponding iteration radius is obtained as the optimal

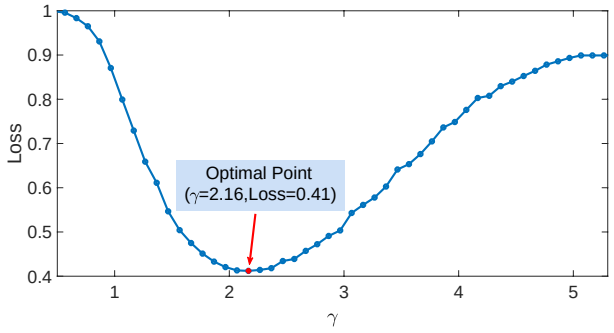

Figure 6: Evolution of loss function against region radius

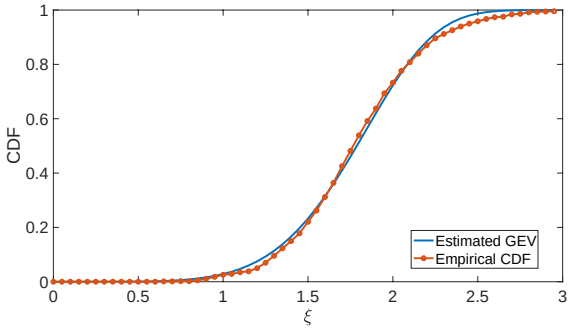

Figure 8: Estimated and empirical CDF curve

one. As an example, the minimum loss and its corresponding radius are marked as a red point in Fig.6.

The second step is the computation of the LMD using the obtained anchors set as referred to in Eq.(4).

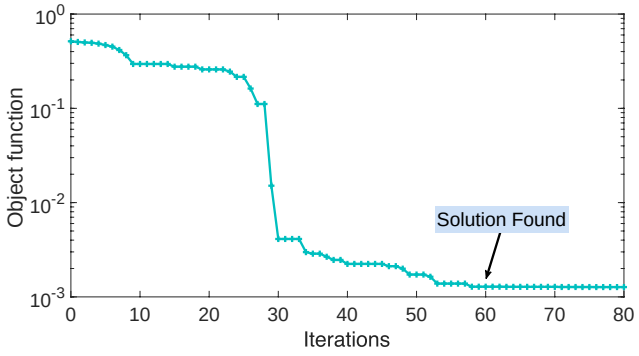

Figure 7: Objective function in the solution searching procedure

Then, the third step is the GEV model estimation. To this end, LMD values for all samples in $\boldsymbol{X}^h$ are used and the Nelder–Mead method is employed to solve the optimization problem of Eq.(13). As shown in Fig.7, the objective value decreases quickly with iterations in the considered procedure. Based on the applied simulation case, after 60 iterations, the solution is found for a low enough considered objective value (close to 10^{-3}).

Fig.8 displays the estimated and the empirical cumulative density function of ξ. The comparison of the two curves indicates that the obtained model has a good approximation for empirical data.

The last step is the selection of a healthy domain margin based on the obtained GEV model. The healthy domain margin ϵ is then determined according to Eq.(14) with the given significance level α.

Based on the elements obtained in this training process (anchor set S, healthy domain margin ϵ, and significance level α), the second part of our proposed procedure (online monitoring) can be computed using new data that can contain faulty ones.

4.2. Online Monitoring

For this seconds part, we consider a new set of samples with a fault occurring in x_1 at $t = 47s$ (the 47^{th} sample) with the fault severity $\delta = 0.02$, as shown in Fig.9.

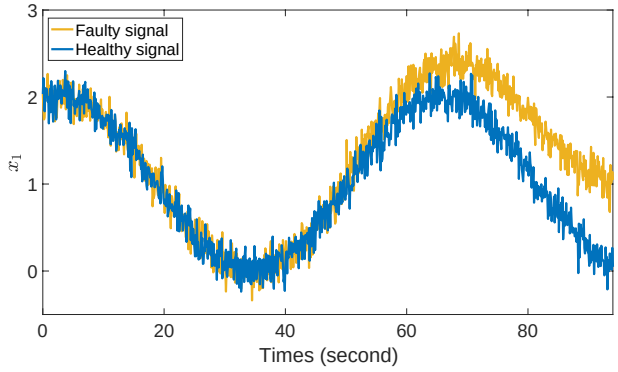

Figure 9: Faulty and healthy x_1 simulated signal

For online monitoring, the LMD value of this new set containing faulty sample $\boldsymbol{x}^f$ is first calculated based on S, as illustrated in Fig.10. Note that the real fault occurrence time ($t = 47s$) is spotted out by the green vertical dashed line.

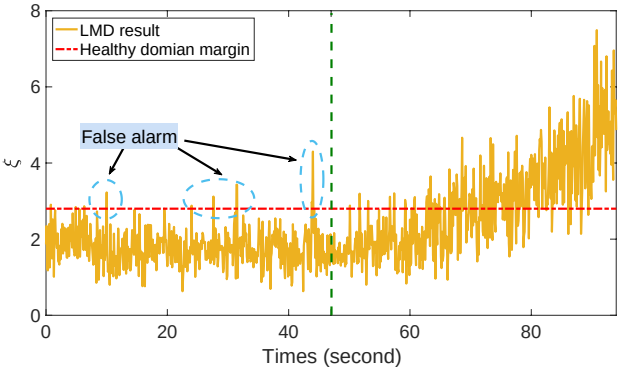

Figure 10: LMD results for simulated signals

Although there are few false alarms in the healthy part of signals, LMD results in this part are basically below the healthy domain margin ϵ marked as a red dashed line in

Fig.10. Nevertheless, in the faulty part (beyond the green vertical dashed line), LMD index increases with time and finally exceed ϵ.

According to this characteristic, LMD is used as the preliminary monitoring index for fault detection. Despite the effectiveness of this LMD index in the long-term for fault detection, missed detections still exist at the beginning of the fault occurrence, as seen in Fig.10, since the fault amplitude is subtle to recognize (incipient fault). Accordingly, it is still a challenge to detect incipient faults in the early stage.

To increase the fault detection accuracy and sensitivity, the Empirical Probability Density Cumulative Sum (EPD-CUSUM) technique, which is reported as a robust and sensitive methodology for incipient fault detection [32], is further applied to the LMD result. However, this technique can not be applied directly in this work because of its strong assumption requiring healthy reference signals with the same time-scale as the faulty one. As a result, the original EPD-CUSUM technique is improved. The main improvements are in the following two aspects:

- First, in the original version, the observed signal is compared with the healthy reference signal to decide the signal's state. Similarly, in this work, LMD index is compared with the healthy domain margin for fault detection. Therefore, the empirical probability density r, representing the probability of the fault occurrence, can recursively be calculated as follows:

$$r(t) = (1-\omega)r(t-1) + \omega I(t) \tag{18}$$

where $r(0) = 0$, $\omega \in (0, 1]$ is a weight factor and

$$I(t) = \begin{cases} 1 & \xi > \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

- Secondly, according to the original setting of the EPD-CUSUM method [32], the probability density under the healthy hypothesis is the constant 0.5, while this value in this work is equal to $1-\alpha$. Thus, by replacing 0.5 with $1-\alpha$, the improved EPD-CUSUM value G is derived as:

$$G(t) = \left\{ \ln \frac{r(t)}{1-\alpha} + G(t-1) \right\}^{+} \tag{20}$$

where $\{\wp\}^{+} = \wp$ for positive values, otherwise it is equal to 0. Without causing ambiguity, the final detection result G will be referred to as the result of LMDA hereafter. When G exceeds the control limit U_G, a fault is detected and the current time is recorded as the fault detection time.

The tuning parameters for ω and U_G mainly depend on noise strength and fault severity. For example, a large ω leads to high sensitivity to incipient faults and strong influence by noise. On the other hand, a high U_G reduces the risk of false alarms but causes considerable detection delay.

The benefit of the EPD-CUSUM technique is twofold. It significantly avoids false alarms and missed detection since only when the LMD result exceeds ϵ continuously, G will increase with time. Besides, this technique improves detection sensitivity because small changes will cumulative with time, yielding a striking value of G. The LMDA result of the simulation data is demonstrated in Fig.11.

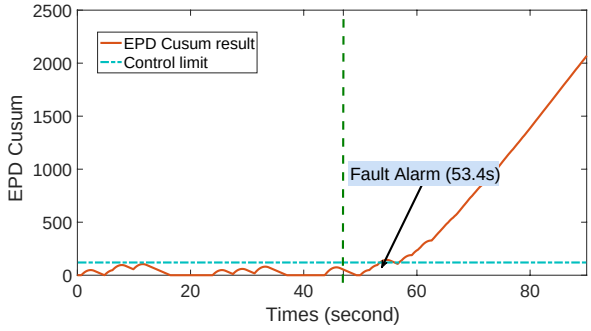

Figure 11: LMDA results for simulated signals using the following parameters: $\omega = 0.3$, $\alpha = 99.9\%$ and $U_G = 110$.

As one can see in this figure, all the LMDA results are below the control limit U_G without false alarm occurring in the healthy part. In the faulty part (beyond the green vertical dashed line) G increase and exceed the control limit at $53.4s$. A delay between the real fault occurrence time and the detected one can then be noticed.

5. Performance analysis

This section investigates the fault detection performance of the proposed method. We study how the proposal will be affected by noise strength and fault severity changes. We propose to probabilistically evaluate the fault detection in these various conditions but also show the effect on the false alarms and the detection delay. Furthermore, the benefit of using the EPD-CUSUM method is highlighted by comparing it with the LMD index D^f and the original cumulative sum technique.

5.1. Analysis criteria

To take into account the effect of fault severity and noise strength simultaneously, the Fault-to-Noise Ratio (FNR) is proposed as the quantitative metric of the fault severity regarding the noise strength [6]. It is defined as:

$$\text{FNR} = 10 \log \frac{p_f}{p_v} \tag{21}$$

where p_f is the fault power and p_v is the noise power. Note that incipient faults are defined for FNR values lower than 5dB. In that range these faults can be partially or totally masked by the noise.

Similarly, the Signal-to-Fault Ratio (SFR) is defined to quantify fault severity regarding the signal power:

$$\text{SFR} = 10 \log \frac{p_s}{p_f} \tag{22}$$

where p_s is the signal power. In the particular case of incipient fault, the SFR values are large.

The above two ratios are related to the common metrics SNR by the following equation:

$$\mathrm{SNR} = \mathrm{SFR} + \mathrm{FNR} \tag{23}$$

The performance evaluation based on samples considers three probabilistic criteria: the detection probability (P_d), the false alarm one (P_{FA}), and the accuracy (ACC) which are defined as:

$$P_d = \frac{\text{No. of samples}\{\text{detected as faulty}|\text{faulty}\}}{\text{No. of }\{\text{All the faulty samples}\}} \tag{24}$$

$$P_{FA} = \frac{\text{No. of samples}\{\text{detected as faulty}|\text{healthy}\}}{\text{No. of }\{\text{All the healthy samples}\}} \tag{25}$$

$$\mathrm{ACC} = \frac{\text{No. of }\{\text{collectly detected samples}\}}{\text{No. of }\{\text{All the samples}\}} \times 100\% \tag{26}$$

Additionally, the Receiving Operating Characteristics (ROC) curve [7] and its corresponding Area Under Curve (AUC) value are also used to globally assess the detection performance in terms of robustness and efficiency.

Moreover, in the early detection context focusing on incipient faults, an overview of the response speed of a fault detection method is also an important criterion. Suppose that a fault occurs at t_0 and the detection time is t_d, the detection delay DD calculated as these two time differences will allow to reflect the method response speed. It is written as:

$$\mathrm{DD} = t_d - t_0 \tag{27}$$

In the following, we further analyze our proposal performance using these criteria.

5.2. Results and discussions

For the following simulations, the signal model given in Eq.(15) in section 4 is used. Signals are observed every 0.1 second for a total duration of 94 seconds, where the number of samples is $K = 940$, and a fault is introduced at $47s$ (i.e. the 470^{th} sample) in the first signal x_1. Besides, Gaussian white noise is added into signals to evaluate the robustness. For the two parts of the proposed procedure, the default settings are summarized below:

- In the training process, the optimal region radius is obtained as $\gamma^{opt} = 2.16$ and then the number of anchors is $\kappa = 96$. So, the GEV model parameters are estimated and the healthy domain margin with the given significance level $\alpha = 99.9\%$ is determined as $\epsilon = 2.79$.
- In the online monitoring procedure, we set $\omega = 0.4$ for the EPD-CUSUM method to adapt to the noise strength with SNR = 20dB.

In this section, using the previously mentioned criteria, we study the training efficiency, the methodology detection sensitivity regarding the fault severity, the robustness of the detection methodology facing noise and the detection time occurrence efficiency.

5.2.1. Training efficiency

The evaluation of training efficiency is to discuss the influence of available training samples' number on the method's performance. In the procedure of healthy domain approximation, two aspects are potentially affected by the number of training samples: the anchors generation and the domain margin selection. In the anchors generation part, the downsampling algorithm is proposed to reduce the repeated samples and extract crucial spatial information of the healthy domain. Therefore, training samples are expected to include as much as possible useful information instead of redundant ones and noises. While, in the domain margin selection, sufficient training samples are necessary to estimate the parameters of the density model. Based on these two considerations, the relation between the training samples' number and the method's performance is then investigated.

Fig.12 exhibits the evolution of accuracy performance along with the change of samples' number. When training samples are insufficient, such as the sample's number is less than 100, viz. the duration of the signals is shorter than 10 seconds, the accuracy performance is poor, lower than 50%. Consequently, when samples' number increases to 300, the result notably increases and almost reaches 70% accuracy. A further increase of the result is obtained when 1000 samples are used for training. However, with the training samples number continually growing from 1000 to 10000, one can notice the sight performance degeneration, which may result from the introduction of a large number of irrelevant samples. As a result, sufficient training samples (> 100) are required, but excess samples (> 5000) will degrades the performance.

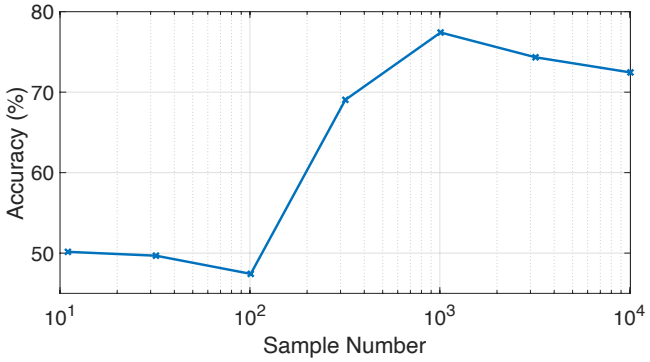

Figure 12: Accuracy results for different numbers of training samples

5.2.2. Detection sensitivity

The second part of the performance analysis focuses on the effect of the fault severity. We consider different values of FNR varying from −10dB to 20dB and keep all the other operating conditions the same as previously defined. In Fig.13, ROC curves of the LMD index D^f for different FNR settings are displayed to show the fault detection capability.

It can be noticed that once FNR decreases from 20dB to −10dB (fault severity decreases), the corresponding AUC

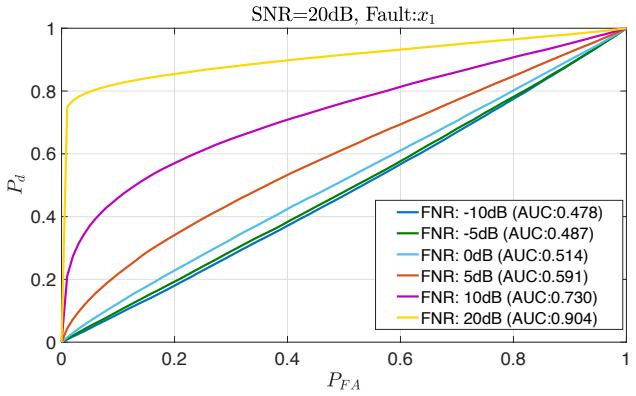

Figure 13: ROC curves of LMD results for different FNR settings

value decreases and reaches a low value close to 0.5 leading to the worst detection performance. When FNR is lower than 5dB, meaning that the fault power is lowest than the triple of the noise power, the fault detection capability of the LMD index is weak.

However, when using the EPD-CUSUM method in the detection procedure, one can see a significant improvement of the detection capability (Fig.14) such as the AUC value for 5dB FNR case increases from 0.591 to 0.708.

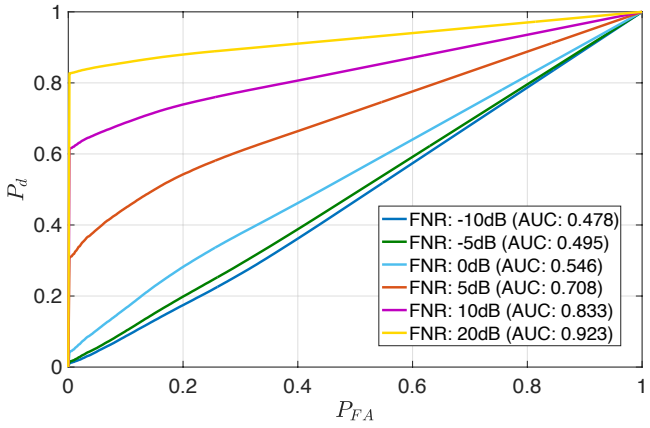

Figure 14: ROC curves of LMDA results for different FNR settings

When we focus on the detection probability performance with a low P_{FA} value, corresponding to a practical condition for fault detection, the improvement results from the EPD-CUSUM method is the most significant. As an example, if we consider $P_{FA} = 0.01$, the detection probabilities of the LMDA results are all larger than that of the LMD index D^f, especially for FNR $= 0$dB, 5dB, 10dB.

To further show the benefit of using EPD-CUSUM, we compare the detection probability of three detection strategies in the case of $P_{FA} = 0.01$: LMD index, LMD with EPD-CUSUM (LMDA), and LMD with the typical CUSUM method [23]. The results are shown in Fig.15 and demonstrate that LMDA notably outperforms the other two strategies, particularly for incipient faults detection (like FNR $<$ 5dB). Nevertheless, the benefit of using the typical CUSUM method is trivial compared to the baseline performance of the LMD index.

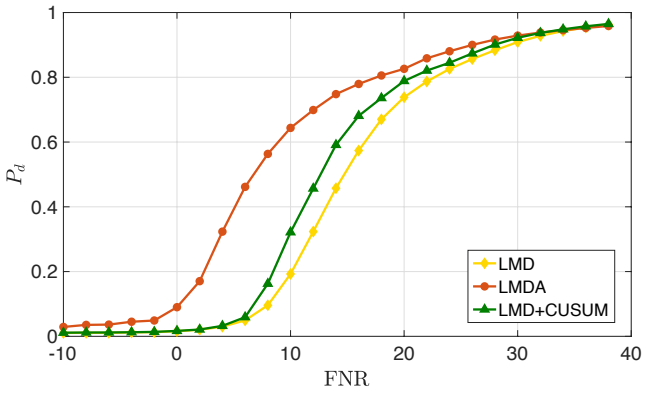

Figure 15: Detection probability performance for different detection strategies, where $P_{FA} = 0.01$

Our proposed methodology offers then a good detection sensitivity and efficiency. For incipient faults that can be widely affected by the noise strength, this sensitivity is widely improved.

5.2.3. Detection robustness

Subsequently, different noise and fault severity levels are taken into account in the simulation to evaluate the performance of our proposal deeply. The goal is then to qualify the robustness of the proposed fault detection methodology regarding noise. Note that Gaussian white noise is added into both the training and testing data. A small part of the noise signal with large values can be regarded as outliers that potentially affect the accuracy of healthy domain approximation.

With SNR varying from 0dB to 50dB and SFR varying from -40dB to 40dB, the resulting AUC values are plotted in Fig.16. The results highlight that both factors affect the performance of the proposed detection method: either the increase of noise power or the decrease of fault severity leads to low detection performance. However, for the common range of noise level, such as from 20dB to 40dB SNR, when the SFR correspondingly increases from 10dB to 30dB, the detection performance of the proposed framework is effective since AUC values in this area are larger than 0.8. In that range the proposed methodology is remaining robust enough.

Although the results indicate that the performance will not dramatically degenerate when training data contains a small number of outliers (20dB to 40dB SNR), it seems that the robustness to the consequently introduced outliers can be further improved by specifically selecting anchors.

5.2.4. Detection time occurrence efficiency

As mentioned in section 4, the fault is time-varying. So, the detection of the fault occurrence time is important to qualify the efficiency of the proposal properly.

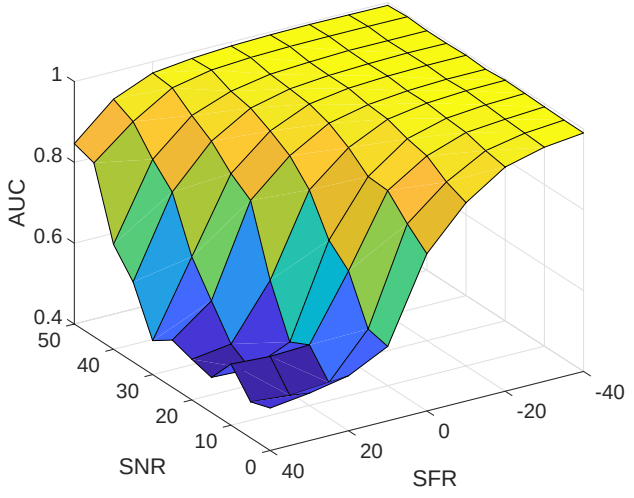

Figure 16: AUC performance for different noise strength and fault severity

As a first study, the noise is settled with SNR = 20dB and we vary the fault severity considering the FNR. The detection delay is evaluated and plotted in Fig.17 for the three previously mentioned techniques (LMD, LMDA, and LMD with CUSUM). For this first study, $P_{FA} = 0.01$ and the detection delay is written in sample number (reminder: the sampling time is 1 sample per 0.1 seconds).

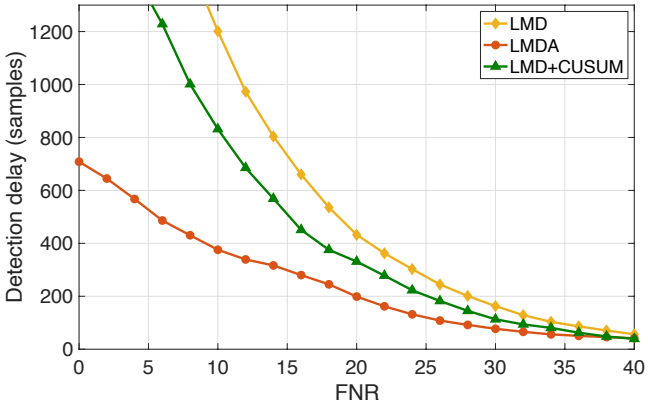

Figure 17: Detection delay performance for different fault severity

It can be noticed that the additional classic cumulative sum procedure with LMD slightly improves the detection delay but our proposed LMDA using EPD-CUSUM is the most efficient one. This latter offers the lowest detection delay all over the FNR range. Nevertheless, the more the fault is incipient (lower FNR) and longer will be the detection delay. As an example, for FNR > 20dB, the detection delay results are lower than 200 samples but if the FNR value decreases, the detection delay exponentially increases.

In a second study, we consider the noise varying in the same SNR range previously mentioned and the fault severity with SFR vary from −40dB to 20dB. The detection delay is evaluated and plotted in Fig.18.

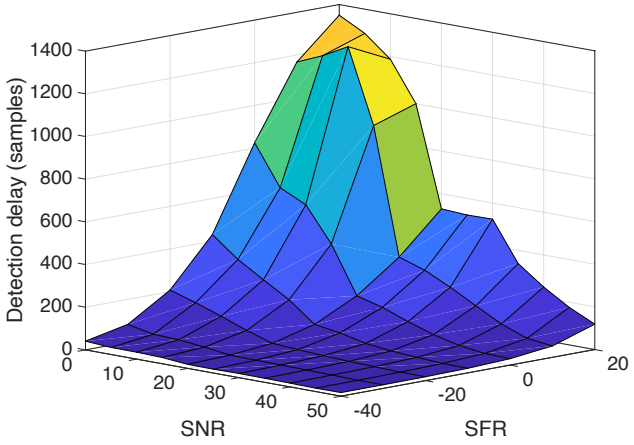

Figure 18: Detection delay performance for different noise strength and fault severity

We can notice that the detection delay remains a low value when SNR is a very high positive value (*i.e* very low noise level) and SFR is a negative very low value (*i.e* very high fault severity). But if the noise widely increases (SNR close to zero) and the fault severity deeply decreases (SFR positive) then the detection delay will exponentially increase and its value will be very high.

Indeed, the detection occurrence is efficient with our proposal. The noise influence has been widely reduced compared to the original LMD proposal but is still remaining and cannot be completely removed. This reduction seems sufficient enough to consider this proposal for an engineering application study.

6. Application to an engineering process

In this section, the Continuous-flow Stirred Tank Reactor (CSTR) process [24, 26] is employed as a case study to validate our proposed methodology. For that purpose, the fault detection performance is evaluated in terms of false alarm probability, detection probability, detection delay, and AUC. The effectiveness of our proposal is then highlighted compared to other state-of-the-art methods known for their efficiency using this process.

6.1. CSTR process description

The continuous-flow stirred reactor (CSTR), related to chemical and environmental engineering, is widely used to simulate incipient faults [24, 26]. It refers to key unit operation variables modeling the behavior of a continuous agitated tank reactor to reach a specified output. The exothermic first-order reaction taking place in the reactor can be modeled as follows:

$$\frac{dC}{dt} = \frac{Q}{V}(C_i - C) - aqC + \nu_1 \tag{28}$$

Table 1
Parameters for CSTR Process

Parameter	Description	Parameter	Description
Q	Inlet flow rate	$\mathcal{T},\mathcal{T}_i,\mathcal{T}_c,\mathcal{T}_{ci}$	Reactor temperature
V	Tank volume	C,C_i	Reactor concentrstion
Q_c	Coolant flow rate	$\mathcal{UA}$	Heat transfer coefficient
V_c	Jacket volume	q_0	pre-exponential factor
ΔH_r	Heat of reaction	$\varphi,\ \varphi_c$	Fluid density
E/R	activation energy	C_p,C_{pc}	Fluid heat capacity
$\nu_{1,2,3}$	process noise		

Table 2
The Description of faulty scenarios

Fault	Description	δ	Fault	Description	δ
1	$a = a_0 e^{-\delta t}$	0.0005	2	$\hbar = \hbar_0 e^{-\delta t}$	0.001
3	fault 1 and 2		4	$C_i = C_{i,0} + \delta t$	0.001
5	$\mathcal{T}_i = T_{i,0} + \delta t$	0.05	6	$\mathcal{T}_{ci} = \mathcal{T}_{ci,0} + \delta t$	0.05
7	$C = C_0 + \delta t$	0.001	8	$\mathcal{T} = \mathcal{T}_0 + \delta t$	0.05
9	$\mathcal{T}_c = \mathcal{T}_{c,0} + \delta t$	0.05	10	$Q_c = Q_{c,0} + \delta t$	-0.1

$$\frac{d\mathcal{T}}{dt} = \frac{Q}{V}(\mathcal{T}_i - \mathcal{T}) - a\frac{\Delta H_r q C}{\varphi C_p} - \hbar\frac{\mathcal{UA}}{\varphi C_p V}(\mathcal{T} - \mathcal{T}_c) + \nu_2 E \quad (29)$$

$$\frac{d\mathcal{T}_c}{dt} = \frac{Q_c}{V_c}(\mathcal{T}_{ci} - \mathcal{T}_c) + \hbar\frac{\mathcal{UA}}{\varphi_c C_{pc} V_c}(\mathcal{T} - \mathcal{T}_c) + \nu_3 \quad (30)$$

where $q = q_0 \exp(\frac{-E}{R\mathcal{T}})$, and the explanation of equations' parameters are given in Table 1. More details about the CSTR process can be obtained in [24]. In the model, the system's input is denoted as $\boldsymbol{u} = [C_i\ \mathcal{T}_i\ \mathcal{T}_{ci}]$ and the output is written as $\boldsymbol{y} = [C\ \mathcal{T}\ \mathcal{T}_c\ Q_c]$.

The process simulation model used in this work is presented by Karl Ezra Pilario *et al.* and available online [24]. It allows to generate the fault-free data and 10 faulty scenarios, where the faulty conditions and their previously defined parameters are summarized in Table 2.

Note that all the faults are time-varying ones with different severities and constructions structure:

- Fault 1 and Fault 2 are multiplicative ones caused by catalyst decay and heat transfer fouling.
- Fault 3 is the combination of the two previous faults (Fault 1 and Fault 2).
- All the others faults are additive ones related to sensor drifts.
- The process noises ν_1, ν_2, ν_3 are additive white Gaussian noise with zero mean and variances $\sigma_1^2 = 0.002$, $\sigma_2^2 = \sigma_3^2 = 2$, respectively, leading to $\mathrm{SNR}_1 = 27\mathrm{dB}$ and $\mathrm{SNR}_2 = \mathrm{SNR}_3 = 47.9\mathrm{dB}$.

6.2. CSTR Fault detection performances

To apply our proposal to this case-study, the process variables are structured as $\boldsymbol{X} = [\boldsymbol{u}\ \boldsymbol{y}]$. We tuned our methodology as follows:

- For the training procedure, we concatenated 10 groups of fault-free data with 1200 samples (i.e 1200 minutes) for each to form the fault-free data matrix $\boldsymbol{X}^h$. Based on $\boldsymbol{X}^h$, the optimal region radius was determined as $\gamma^{opt} = 1.448$ and then 224 anchors were generated. The healthy domain margin was obtained as $\epsilon = 2.7$ with the given significance level $\alpha = 99.9\%$.
- To evaluate the online monitoring performance, 250 groups of data for each faulty scenario were generated, where each group contained 1200 samples, and faults were introduced at $t = 200$ minutes (the 200^{th} sample). For the EPD-CUSUM calculation, the tuning parameters were set as $\omega = 0.4$ and $U_G = 110$.

The obtained detection results for the multiplicative type faults (Fault $1-3$) are illustrated in Fig.19, and those for additive type faults (Fault $4-10$) are displayed in Fig.20. On the left side of these figures, LMD results are shown in yellow, while the margin is marked as a horizontal red dashed line. On the right side of the figures, LMDA results are in orange, and the control limit for the decision-making is marked as a horizontal cyan dashed line.

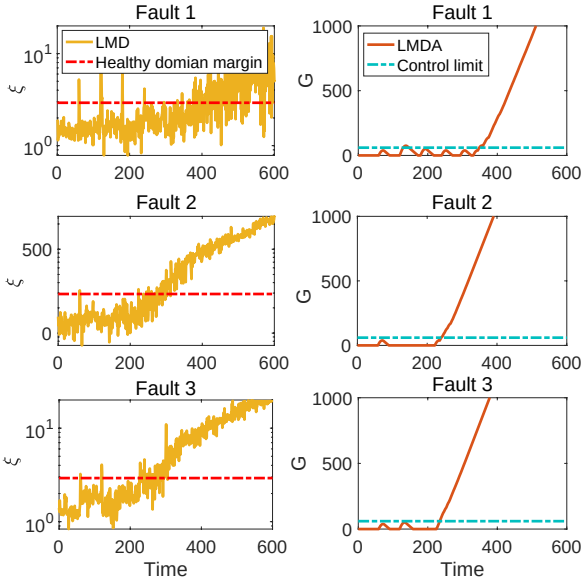

Figure 19: LMD and LMDA results for Faults $1-3$.

For the two types of faults, LMD and LMDA results increase with time after faults occurrence. Although false alarms exist in the healthy part of the LMD results, the right figures highlight that LMDA results can significantly alleviate the false alarm issue by turning incidental impulse-like detected results into slight fluctuation. Additionally, using the EPD-CUSUM technique allows to summarize those small continuous deviations in LMD results, producing growing results with time. This helps to have a better detection for incipient faults and shorten the detection delay. Moreover, due to the cumulative nature of the EPD-CUSUM method, missed detection can be remarkably avoided and missed detection can be performed.

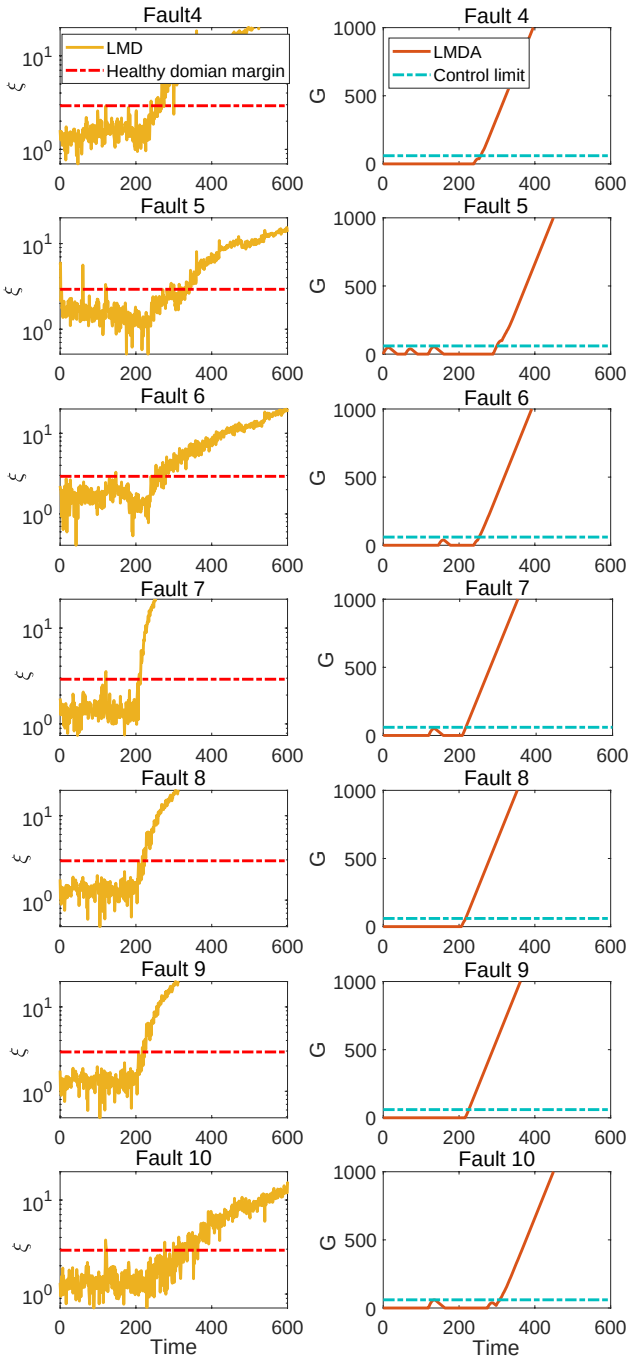

Figure 20: LMD and LMDA results for Faults 4 – 10

Table 3
Fault diagnosis performance for CSTR

Fault	GCCA		CVA		CVDA	LMDA
	T^2_{r1}	T^2_{r2}	T^2	SPE	D	G
1	11.05	10.54	4.18	2.74	**1.46**	1.50
	0.07	0.13	0.57	0.57	0.89	1.39
	48.22	50.88	70.88	82.35	**91.05**	86.46
	0.895	0.943	0.954	0.973	**0.981**	0.976
2	9.47	2.95	4.82	2.63	1.59	**1.11**
	0.05	0.12	0.59	0.64	0.93	1.52
	39.95	83.19	69.74	83.51	89.98	**92.49**
	0.968	0.977	0.912	0.971	0.979	**0.982**
3	8.33	2.93	3.39	2.42	1.54	**0.93**
	0.04	0.09	0.46	0.53	0.75	1.97
	53.97	83.51	77.43	85.04	90.32	**93.54**
	0.973	0.979	0.958	0.977	**0.983**	**0.983**
4	2.33	2.30	2.00	2.35	1.44	**0.74**
	0.07	0.11	0.52	0.67	0.83	1.03
	86.46	86.73	88.52	85.55	91.38	**95.11**
	0.986	0.984	0.972	0.978	0.982	**0.987**
5	3.66	9.13	2.82	4.41	2.50	**1.30**
	0.05	0.08	0.33	0.50	0.57	1.03
	78.23	47.22	82.90	71.74	84.00	**90.84**
	0.970	0.939	0.960	0.963	0.971	**0.974**
6	8.73	3.07	3.90	2.90	2.19	**1.04**
	0.08	0.12	0.57	0.54	0.89	0.95
	48.36	82.47	76.66	81.92	86.12	**92.73**
	0.979	0.973	0.964	0.973	0.977	**0.981**
7	0.41	0.38	0.36	0.23	**0.15**	0.17
	0.07	0.11	0.55	0.52	0.83	1.11
	97.78	97.93	98.04	98.69	**99.18**	98.93
	0.995	**0.997**	0.992	0.992	0.994	0.996
8	2.29	0.88	0.87	0.53	**0.37**	**0.37**
	0.07	0.11	0.51	0.64	0.77	1.60
	87.19	95.24	95.15	96.93	**97.92**	97.70
	0.991	0.993	0.989	0.988	0.992	**0.994**
9	6.62	0.80	1.48	0.80	0.53	**0.34**
	0.09	0.14	0.66	0.66	1.12	1.13
	62.27	95.62	91.49	95.42	96.95	**97.87**
	0.992	**0.994**	0.970	0.988	0.992	**0.994**
10	10.51	4.35	5.22	3.38	2.37	**1.31**
	0.06	0.09	0.34	0.50	0.60	1.13
	37.9	0 75.26	68.27	78.89	84.99	**90.39**
	0.962	0.969	0.937	0.968	0.974	**0.982**

* For each fault: First row is the Detection Delays (hours); Second row is the False alarm probability (%); Third row is the Detection probability (%); Fourth row is the AUC value

Furthermore, to highlight the benefit of our proposal, LMDA (with EPD-CUSUM) results are compared with other fair literature methods reported as effective to non-Gaussian distributed data and developed for incipient fault detection. We have selected the Generalized Canonical Correlation Analysis (GCCA) [4], Canonical Variate Analysis (CVA) [21], and Canonical Variate Dissimilarity Analysis (CVDA) method [24]. For each of them, the results in terms of detection delay, false alarm probability, detection probability, and AUC value are collected and summarized in Table 3. Note that due to the large time constant in the CSTR process, the detection delay is presented in hours. The false alarm and detection probabilities are given as ratios and presented in percentages. The best values overall the reported techniques are highlighted in the table with bold font.

In terms of detection delay, our proposal is efficient to detect faults with less than 1.5 of hours detection delay (no more than 90 additional sample points). This outperforms the other three approaches for most cases except for Fault 1 and 7. GCCA is not sensitive enough to incipient fault detection as its detection delay for both indexes is quite large. Although CVDA and CVA seem effective in the early detection of incipient faults, they are not as efficient as the LMDA methodology that we propose.

Nevertheless, in terms of false alarms, the GCCA technique is the most reliable among the four methods. As for our proposal, the false alarm probability is not perfect but acceptable with small rate values always lower than 1.97%.

Even with the not perfect false alarm probability, our proposal has the best performance in terms of detection probability for most of the fault conditions. Only CVDA in the case of Fault 1, 7 and, 8 offers slightly better rate values (less than 0.25% except for Fault 1 where the improvement is close to 4.6%).

Using the AUC criterion for representing the global performance evaluation, LMDA offers the largest value for most cases except in the case of Fault 1 and 7. The difference with the best results shown GCCA is only 0.001. This indicates that the results that we obtain are fully acceptable. Our proposal can then be viewed with the best overall performances for the CSTR case study.

Our solution offers the best trade-off performances compared to the other techniques. Fault 1 has particularly more tricky results. The impact of the noise seems more affecting this incipient fault detection.

7. Conclusion

This paper propose a specific incipient fault detection methodology for time-varying faults in a noisy environment. It is based on a procedure composed of two main parts: the training process and the online monitoring.

In the training process, we propose a specific unsupervised healthy domain approximation method using the Mahalanobis distance in a particular way. It is composed of an optimized down-sampling algorithm and a special margin selection approach. Compared with PCA-based Hotelling's T^2 and SPE (Q) statistics, the proposed healthy domain approximation method can effectively handle non-elliptically distributed data.

Based on the approximated healthy domain characteristics, the online monitoring process is derived and the local Mahalanobis distance is defined and served as a preliminary monitoring index for fault detection. Subsequently, we derive this sensitive incipient fault detection framework combining Local Mahalanobis Distance with the improved Empirical Probability Density Cumulative Sum method.

The performance analysis based on the simulation data shows that the proposed fault detection method is effective for data without any distribution-type assumption. The proposal is robust to noise influence and achieves pronounced performances in incipient fault detection (high sensitivity to fault). It outperforms other classical methods. The comparison result also demonstrates the potential benefit of combining the improved EPD-CUSUM approach with other typical process monitoring techniques to improve incipient fault detection capability.

We consider the engineering case study of the CSTR process to validate our proposal. Compared to other well-tuned and efficient state-of-the-art methods, the detection capability of our proposal (LMDA) outperforms these techniques in terms of false alarm probability, detection probability, detection delay, and AUC.

For future works, the characterisation (estimation, isolation, ...) of the detected fault using this approach should be investigated. Moreover, the wide limits and robustness improvement to outliers in the training data will be further considered and evaluated.

Acknowledgments

The authors would like to thank China Scholarship Council for funding.

A. Appendix

Introduction to extreme value statistics

Theorem 1. *Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed (i.i.d) random variables with cumulative distribution function $F(X)$ and consider their maximum value M_n of n samples, i.e. $M_n = \max\{X_1, \cdots, X_n\}$. The induced distribution of M_n can only take one of three forms Gumbel, Weibull, or Frechet with ρ, β, τ respectively denoted as the location, scale and shape parameters:*

- *Gumbel*

$$\overline{F}_G(x,\rho,\beta) = \exp\left[-\exp\left(-\frac{x-\rho}{\beta}\right)\right] \qquad x\in\mathbb{R}, \quad \beta>0 \tag{31}$$

- *Weibull*

$$\overline{F}_W(x,\rho,\beta,\tau) = \begin{cases} 1 & x \geq \rho \\ \exp\left[-\left(\frac{\rho-x}{\beta}\right)^{\tau}\right] & \text{otherwise} \end{cases} \tag{32}$$

- *Frechet*

$$\overline{F}_F(x,\rho,\beta,\tau) = \begin{cases} \exp\left[-\left(\frac{\beta}{x-\rho}\right)^{\tau}\right] & x \geq \rho \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

Let $m_n = \min\{X_1, \cdots, X_n\}$. Similarly, the induced distribution of m_n can be one of three:

- *Gumbel*

$$\underline{F}_G(x,\rho,\beta) = 1-\exp\left[-\exp\left(\frac{x-\rho}{\beta}\right)\right] \qquad x\in\mathbb{R}, \quad \beta>0 \tag{34}$$

- *Weibull*

$$\underline{F}_W(x,\rho,\beta,\tau) = \begin{cases} 0 & x \leq \rho \\ 1-\exp\left[-\left(\frac{x-\rho}{\beta}\right)^{\tau}\right] & x > \rho \end{cases} \tag{35}$$

- *Frechet*

$$\underline{F}_F(x,\rho,\beta,\tau) = \begin{cases} 1-\exp\left[-\left(\frac{\rho}{\rho-x}\right)^{\tau}\right] & x \le \rho \\ 1 & \text{otherwise} \end{cases} \tag{36}$$

Eqs.(31)–(33) can be unified into a simple form for maxima

$$\overline{\Phi}(x,\rho,\beta,\tau) = \exp\left\{-\left[1+\tau\left(\frac{x-\rho}{\beta}\right)\right]^{-1/\tau}\right\} \tag{37}$$
$$-\beta-\tau(x-\rho)\le 0, \qquad \beta>0$$

Eqs.(34)–(36) can be unified into a simple form for minima

$$\underline{\Phi}(x,\rho,\beta,\tau) = 1-\exp\left\{-\left[1+\tau\left(\frac{\rho-x}{\beta}\right)\right]^{-1/\tau}\right\} \tag{38}$$
$$-\beta-\tau(\rho-x)\le 0, \qquad \beta>0$$

For more details about extreme value statistics, readers can refer to [13, 2].

References

[1] Arunthavanathan, R., Khan, F., Ahmed, S., Imtiaz, S., Rusli, R., 2020. Fault detection and diagnosis in process system using artificial intelligence-based cognitive technique. Computers & Chemical Engineering 134, 106697.

[2] Balakrishnan, N., Nevzorov, V.B., 2003. A primer on statistical distributions. Wiley & Sons, Hoboken, New Jersey, USA.

[3] Basseville, M., 1989. Distance measures for signal processing and pattern recognition. Signal processing 18, 349–369.

[4] Chen, Z., Ding, S.X., Peng, T., Yang, C., Gui, W., 2017. Fault detection for non-gaussian processes using generalized canonical correlation analysis and randomized algorithms. IEEE Transactions on Industrial Electronics 65, 1559–1567.

[5] Chen, Z., Mauricio, A., Li, W., Gryllias, K., 2020. A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. Mechanical Systems and Signal Processing 140, 106683.

[6] Delpha, C., Diallo, D., 2015. Incipient fault detection and diagnosis: a hidden information detection problem, in: 2015 IEEE 24th International Symposium on Industrial Electronics (ISIE), IEEE. pp. 837–842.

[7] Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognition Letters 27, 861 – 874.

[8] Golosnoy, V., Ragulin, S., Schmid, W., 2009. Multivariate cusum chart: properties and enhancements. AStA Advances in Statistical Analysis 93, 263–279.

[9] Harmouche, J., Delpha, C., Diallo, D., 2015. Incipient fault detection and diagnosis based on Kullback–Leibler divergence using principal component analysis: Part II. Signal Processing 109, 334–344.

[10] Harmouche, J., Delpha, C., Diallo, D., 2016. Incipient fault amplitude estimation using KL divergence with a probabilistic approach. Signal Processing 120, 1–7.

[11] Isermann, R., 2005. Model-based fault-detection and diagnosis–status and applications. Annual Reviews in control 29, 71–85.

[12] Khan, S.S., Madden, M.G., 2009. A survey of recent trends in one class classification, in: Irish conference on artificial intelligence and cognitive science, Springer. pp. 188–197.

[13] Kotz, S., Saralees, N., 2000. Extreme Value Distributions: Theory and Applications. London Imperial College Press.

[14] Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E., 1998. Convergence properties of the Nelder–Mead simplex method in low dimensions. SIAM Journal on optimization 9, 112–147.

[15] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., Nandi, A.K., 2020. Applications of machine learning to machine fault diagnosis: A review and roadmap. Mechanical Systems and Signal Processing 138, 106587.

[16] Liu, J., Li, Y.F., Zio, E., 2017. A SVM framework for fault detection of the braking system in a high speed train. Mechanical Systems and Signal Processing 87, 401–409.

[17] Lorden, G., 1971. Procedures for reacting to a change in distribution. The Annals of Mathematical Statistics 42, 1897–1908.

[18] Lowry, C.A., Woodall, W.H., Champ, C.W., Rigdon, S.E., 1992. A multivariate exponentially weighted moving average control chart. Technometrics 34, 46–53.

[19] Mansouri, M., Nounou, M.N., Nounou, H.N., 2017. Multiscale kernel PLS-based exponentially weighted-GLRT and its application to fault detection. IEEE Transactions on Emerging Topics in Computational Intelligence 3, 49–58.

[20] Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. The computer journal 7, 308–313.

[21] Odiowei, P.E.P., Cao, Y., 2009. Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. IEEE Transactions on Industrial Informatics 6, 36–45.

[22] Page, E.S., 1954. Continuous inspection schemes. Biometrika 41, 100–115.

[23] Page, E.S., 1961. Cumulative sum charts. Technometrics 3, 1–9.

[24] Pilario, K.E.S., Cao, Y., 2018. Canonical variate dissimilarity analysis for process incipient fault detection. IEEE Transactions on Industrial Informatics 14, 5308–5315.

[25] Roberts, S.W., 2000. Control chart tests based on geometric moving averages. Technometrics 42, 97–101.

[26] Shang, L., Yan, Z., Li, J., Qiu, A., Zhang, H., 2020. Canonical residual based incipient fault detection method for industrial process, in: 2020 Chinese Control And Decision Conference (CCDC), IEEE. pp. 987–992.

[27] Shewhart, W.A., 1931. Economic control of quality of manufactured product. Macmillan And Co Ltd, London.

[28] Spendley, W., Hext, G.R., Himsworth, F.R., 1962. Sequential application of simplex designs in optimisation and evolutionary operation. Technometrics 4, 441–461.

[29] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., Yin, K., 2003. A review of process fault detection and diagnosis: Part III: Process history based methods. Computers & chemical engineering 27, 327–346.

[30] Wang, Y., Pan, Z., Yuan, X., Yang, C., Gui, W., 2020. A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network. ISA transactions 96, 457–467.

[31] Wen, L., Gao, L., Li, X., 2017. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. IEEE Transactions on systems, man, and cybernetics: systems 49, 136–144.

[32] Yang, J., Delpha, C., 2020a. Empirical probability density cumulative sum for incipient fault detection, in: 2020 Prognostics and Health Management Conference (PHM-Besançon), IEEE. pp. 187–193.

[33] Yang, J., Delpha, C., 2020b. Open-circuit fault diagnosis for interleaved DC-DC converters, in: IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, IEEE. pp. 3982–3987.

[34] Yin, S., Ding, S.X., Haghani, A., Hao, H., Zhang, P., 2012. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. Journal of process control 22, 1567–1581.

[35] Yin, S., Ding, S.X., Xie, X., Luo, H., 2014. A review on basic data-driven approaches for industrial process monitoring. IEEE Transactions on Industrial Electronics 61, 6418–6428.

[36] Youssef, A., Delpha, C., Diallo, D., 2016. An optimal fault detection threshold for early detection using Kullback–Leibler divergence for unknown distribution data. Signal Processing 120, 266–279.

[37] Ypma, A., Duin, R.P.W., 1998. Support objects for domain approx-

imation, in: International Conference on Artificial Neural Networks, Springer. pp. 719–724.

[38] Zeng, J., Kruger, U., Geluk, J., Wang, X., Xie, L., 2014. Detecting abnormal situations using the Kullback–Leibler divergence. Automatica 50, 2777–2786.

[39] Zhang, M., Jiang, Z., Feng, K., 2017. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. Mechanical Systems and Signal Processing 93, 460–493.

[40] Zhang, X., Delpha, C., Diallo, D., 2020. Incipient fault detection and estimation based on Jensen–Shannon divergence in a data-driven approach. Signal Processing 169, 107410.